# Estimation of the control parameter from symbolic sequences: Unimodal maps with variable critical point

David Arroyo,[1,a] Gonzalo Alvarez,[1] and José María Amigó[2]
[1]*Instituto de Física Aplicada, Consejo Superior de Investigaciones Científicas, Serrano 144, 28006 Madrid, Spain*
[2]*Centro de Investigación Operativa, Universidad Miguel Hernández, Avda. de la Universidad s/n, 03202 Elche, Spain*

(Received 3 December 2008; accepted 26 May 2009; published online 26 June 2009)

The work described in this paper can be interpreted as an application of the order patterns of symbolic dynamics when dealing with unimodal maps. Specifically, it is shown how Gray codes can be used to estimate the probability distribution functions (PDFs) of the order patterns of unimodal maps whose dynamics is controlled by an external parameter. Furthermore, these PDFs depend on the value of the external parameter, which eventually provides a handle to estimate the parameter value from symbolic sequences (in form of Gray codes), even when the critical point depends on the parameter. © *2009 American Institute of Physics*. [DOI: 10.1063/1.3155072]

**In this paper, the order patterns of unimodal maps are studied. It is shown how to construct order patterns of unimodal maps from their symbolic dynamics with respect to the partition of the state space introduced by the critical point. Finally, it is shown that for a subclass of parametric unimodal maps, the study of those order patterns allows to estimate the parameter of the map that has generated the symbolic sequence.**

## I. INTRODUCTION

Sarkovskii's theorem shows that order and dynamics are intertwined in one-dimensional intervals. It is therefore not surprising that the study of the ordinal structure of deterministic time series gives valuable information on the underlying dynamical system. This work focuses on the reconstruction of the so-called order patterns of certain unimodal maps from "coarse-grained" orbits in form of 0-1 sequences: 0 if the corresponding iterate lies to the left of the critical point and 1 otherwise. Such binary sequences will be called Gray codes. The relationship between the Gray codes of parametric unimodal maps and the value of the parameter that controls a particular dynamic was shown in Refs. 1–3.

Other important tool for the understanding of one-dimensional dynamical systems is the study of their order patterns.[4] Indeed, order patterns allow to distinguish chaos from white noise[5] and can provide useful information on the parameter or parameters controlling the dynamics of chaotic systems. The main goal of this paper is to estimate the control parameter of unimodal maps by means of their order patterns alone, even when the exact values of their orbits are not accessible but only the corresponding Gray codes. To our knowledge, this is a novel technique to recover control parameters of a dynamic based on symbolic sequences. The comparison with other strategies in the estimation of mapping parameters generating an orbit remains to be done and will be the subject of future research.

Possible applications include the cryptanalysis of chaotic stream cipher, which is a topic currently under investigation. More generally, real time series (such as experimental observations or numerical simulations) are always coarse-grained versions of the actual values on account of the finite precision of observational devices and computer arithmetic. This being the case, our paper touches on a basic and difficult problem if in a simplified setting.

The rest of the paper is organized as follows. First of all, the general framework is set in Sec. II. In Sec. III, the concept of order pattern is introduced and its dependence on the control parameter is analyzed for the logistic and the skew tent maps. Section IV summarizes the theory on Gray codes. How the order patterns of unimodal maps are obtained using Gray codes is explained in Sec. V; its application to control parameter estimation is explained in Sec. VI. The results presented in this paper are recapitulated in Sec. VII, where some final comments are also included.

## II. SCENARIO

The work described in this paper focuses on the class of *unimodal maps*, hereafter denoted as $\mathcal{F}$. A map $f: I \to I$, where $I=[a,b] \subset \mathbb{R}$, $a<b$, is unimodal if it is continuous, has a single turning point (usually called the critical point) $x_c$ in $I$, and is monotone increasing on the left of $x_c$ and decreasing on the right. The class $\mathcal{F}$ includes maps defined in a parametric way, say, $f_\lambda(x)=\varphi(\lambda,x)$, where $x \in I=[a,b]$, $\lambda \in J \subset \mathbb{R}$ is called the *control parameter*, and $\varphi$ is a map on $I \times J$. Two different situations are considered in this paper:

(1) The control parameter determines the maximum value of the map. In this case, the parametric function $f_\lambda$ is given by

[a]Electronic mail: david.arroyo@iec.csic.es.

 

TABLE I. Order patterns of length 4.

| No. | Order pattern | No. | Order pattern | No. | Order pattern | No. | Order pattern |
|---|---|---|---|---|---|---|---|
| 0 | [0,1,2,3] | 1 | [0,1,3,2] | 2 | [0,3,1,2] | 3 | [3,0,1,2] |
| 4 | [3,0,2,1] | 5 | [0,3,2,1] | 6 | [0,2,3,1] | 7 | [0,2,1,3] |
| 8 | [2,0,1,3] | 9 | [2,0,3,1] | 10 | [2,3,0,1] | 11 | [3,2,0,1] |
| 12 | [3,2,1,0] | 13 | [2,3,1,0] | 14 | [2,1,3,0] | 15 | [2,1,0,3] |
| 16 | [1,2,0,3] | 17 | [1,2,3,0] | 18 | [1,3,2,0] | 19 | [3,1,2,0] |
| 20 | [3,1,0,2] | 21 | [1,3,0,2] | 22 | [1,0,3,2] | 23 | [1,0,2,3] |

$$f_\lambda(x) = \lambda F(x), \tag{1}$$

where $F \in \mathcal{F}$ and $F(x_c)=F_{\max}$. The subclass of maps $f_\lambda \in \mathcal{F}$ complying with this description will be denoted by $\mathcal{F}_1$.

(2) The control parameter is the value of the critical point, i.e., $x_c=\lambda$. This leads to a new subclass of maps $\mathcal{F}_2$.

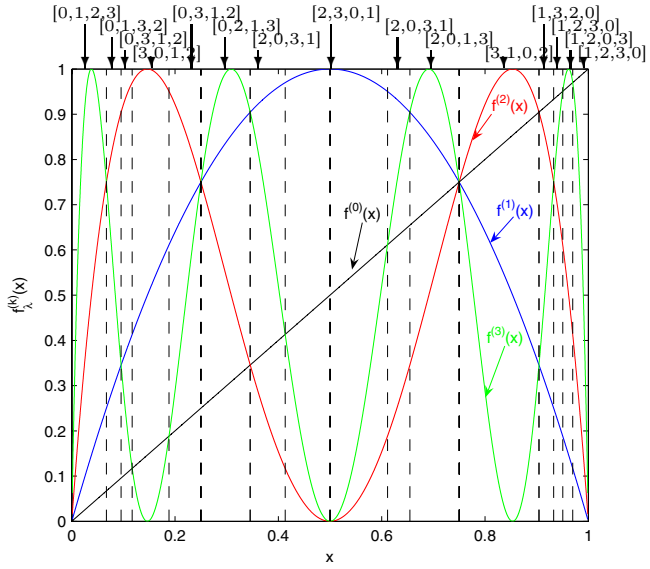


FIG. 1. (Color online) $f_\lambda^{(k)}(x)$ for $k=0,1,2,3$ and the corresponding order patterns of length 4 for the logistic map when $\lambda=4$.

## III. ORDER PATTERNS

Given a closed interval $I\subset\mathbb{R}$ and a map $f:I\to I$, the *orbit* of (the initial condition) $x\in I$ is defined as the set $\mathcal{O}_f(x)=\{f^n(x):n\in\mathbb{N}_0\}$, where $\mathbb{N}_0=\{0\}\cup\mathbb{N}=\{0,1,\ldots\}$, $f^0(x)=x$, and $f^n(x)=f(f^{n-1}(x))$. Orbits are used to define *order L-patterns* (or order patterns *of length L*), which are permutations of the elements $\{0,1,\ldots,L-1\}$, $L\geq 2$. We write $\pi=[\pi_0,\pi_1,\ldots,\pi_{L-1}]$ for the permutation $\pi_0\mapsto 0,\ldots,\pi_{L-1}\mapsto L-1$.

**Definition 1:** Order pattern. The point $x\in I$ is said to define (or realize) the order $L$-pattern $\pi=\pi(x)=[\pi_0,\pi_1,\ldots,\pi_{L-1}]$ if

$$f^{\pi_0}(x) < f^{\pi_1}(x) < \cdots < f^{\pi_{L-1}}(x). \tag{2}$$

Alternatively, $x$ is said to be of type $\pi$. The set of all possible order patterns of length $L$ is denoted by $\mathcal{S}_L$.

For further reference, it is convenient to assign an integer number to each order pattern. This can be made, for instance, by means of the Trotter–Johnson algorithm.[6] The order patterns of length 4 along with their "ordering numbers" are shown in Table I.

As emphasized in Ref. 7 there always exist order $L$-patterns with sufficiently large $L$ that are not realized in any orbit of $f\in\mathcal{F}$. These order patterns are called *forbidden patterns*, whereas the rest of order patterns are called *allowed patterns*. In general, if $f_\lambda$ is a family of self-maps of the closed interval $I\subset\mathbb{R}$ parametrized by $\lambda\in J\subset\mathbb{R}$ (as it occurs for $f_\lambda\in\mathcal{F}_1,\mathcal{F}_2$) and the set $P_\pi$ is defined as

$$P_\pi=\{x\in I: x \text{ is of type } \pi\}, \tag{3}$$

where $\pi\in\mathcal{S}_L$, then $P_\pi$ depends on $f_\lambda$ and, consequently, on $\lambda$. Moreover, we will assume that $f_\lambda$ is ergodic for $J\subset\mathbb{R}$ so as the orbits of $f_\lambda$ can be used to build up statistics independently from the value of the initial condition. Remember that $f_\lambda:I\to I$ is said to be *ergodic* with respect to the invariant measure $\mu$ if the only invariant sets are the empty set and the full state space $I$, except possibly for a $\mu$-null set. According to Birkhoff's ergodic theorem,[8] if $f_\lambda$ is ergodic with respect to the invariant measure $\mu$, then the orbit of $x\in I$ visits the set $P_\pi$ with relative frequency $\mu(P_\pi)$ for almost all $x$ with respect to $\mu$. As a result, it is possible to study the dependence of $P_\pi$ on $\lambda$ by counting and normalizing the occurrences of $\pi$ in sliding windows of width $L$ along $\mathcal{O}_{f_\lambda}(x)$, $x$ being a "typical" initial condition. In Secs. III A and III B this is done experimentally with the logistic map (as representative of $\mathcal{F}_1$) and with the skew tent map (as representative of $\mathcal{F}_2$). Since we are primarily interested in the relation between the probabilities $\mu(P_\pi)$ (or relative frequencies) of order patterns $\pi\in\mathcal{S}_L$ and the control parameter $\lambda$ of the map considered, we will refer to it as the $\lambda$-distribution function (in short, $\lambda$-DF) of $\pi$, since they are related to the probability distribution functions (we fix $\pi$ instead of fixing $\lambda$).

### A. Order patterns for the logistic map

The logistic map defined as

$$f_\lambda(x) = \lambda x(1-x) \tag{4}$$

for $x\in[0,1]$ and $\lambda\in[1,4]$ belongs to $\mathcal{F}_1$. The logistic map with $\lambda=4$ was studied in Refs. 7 and 9 from the ordinal point of view. In Fig. 1 the allowed order four patterns for the logistic map with $\lambda=4$ are shown. For this value of the control parameter there exist twelve allowed order patterns. However, the main goal of this paper is to analyze the relationship between the control parameter of maps in $\mathcal{F}_1$ or $\mathcal{F}_2$ and their order patterns, which calls for the distributions of allowed patterns for different values of $\lambda$. Figure 2 depicts the relative frequencies of each order four patterns for $\lambda\in[3.7,4]$, the patterns being labeled as in Table I. To be more specific, for every $\lambda$, a sufficiently long orbit was generated, the occurrences of the different order patterns were counted using a sliding window of width 4, and finally the counts obtained were normalized by the number of windows. These results are estimates of the probabilities for the corresponding order patterns to occur. Let us point out that since the physical invariant measure of the logistic map is only

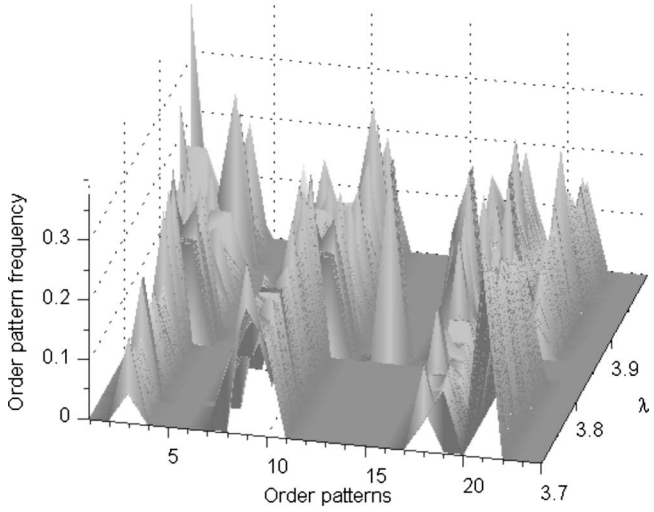


FIG. 2. Relative frequency of the order patterns realized by the logistic map when $L=4$ and $\lambda \in [3.7,4]$.

known for $\lambda=4$, the numerical estimation of those probabilities is the most we can hope for. More importantly for us, we conclude from Fig. 2 that it is very difficult to infer the value of $\lambda \in [3.7,4]$ from the $\lambda$-DF of order patterns of length 4.

### B. Order patterns for the skew tent map

The skew tent map given by

$$f_\lambda(x) = \begin{cases} x/\lambda & \text{if } 0 \le x < \lambda, \\ (1-x)/(1-\lambda) & \text{if } \lambda \le x \le 1 \end{cases} \tag{5}$$

for $x \in [0,1]$ and $\lambda \in (0,1)$ belongs to the subclass $\mathcal{F}_2$, comprised of those maps of $\mathcal{F}$ parametrized by the critical point. Furthermore, for the skew tent map $f_\lambda$, the maximum value $f_\lambda(x_c)=f_\lambda(\lambda)=1$ is independent of $\lambda$ (see Fig. 3). Contrary to the logistic map, the skew tent map does possess a known ergodic invariant measure for all $\lambda \in (0,1)$, namely, the Lebesgue measure on $[0,1]$ (see Ref. 10). Hence, if $P_\pi$ is given by Eq. (3) with $I=[0,1]$, the relative frequency of the order pattern $\pi$ in a typical orbit of the skew tent map coincides with the Lebesgue measure of $P_\pi$, which can be determined analytically. The easiest case corresponds to the order pattern $\pi=[0,1,\ldots,L-1]$, since then $P_\pi$ is an open interval whose left end point is 0 and whose right end point is the leftmost intersection between $f_\lambda^{L-1}$ and $f_\lambda^{L-2}$. The relative frequencies of the order patterns of length 4, numbered according to Table I, are depicted in Fig. 4. In particular, the length of the interval $P_{[0,1,2,3]}=:(0,\phi_4(\lambda))$ is determined by the first intersection between $f_\lambda^{(2)}(x)$ and $f_\lambda^{(3)}(x)$,

$$\phi_4(\lambda) = \frac{\lambda^2}{2-\lambda}. \tag{6}$$

Therefore, the $\lambda$-DF of $\pi=[0,1,2,3]$ (pattern 0) is given by $\phi_4(\lambda)$; see Fig. 5(a) for the graphical representation of $\phi_4(\lambda)$. The fact that the function $\phi_4(\lambda)$ is bijective entails the possibility of estimating $\lambda$ via the relative frequency of the order pattern [0,1,2,3].

Up to this point it has been assumed that the orbits of the various maps considered were accessible. From a more practical point of view, it is also relevant to know whether order patterns can be still determined using less information about the orbits. This is the case, for instance, when dealing with the symbolic dynamics associated with a generating partition of the state space. In particular, the orbits of maps of $\mathcal{F}$ can be transformed into binary sequences by the procedure described in Ref. 1. In Sec. IV it is explained how to build order patterns from those binary sequences.

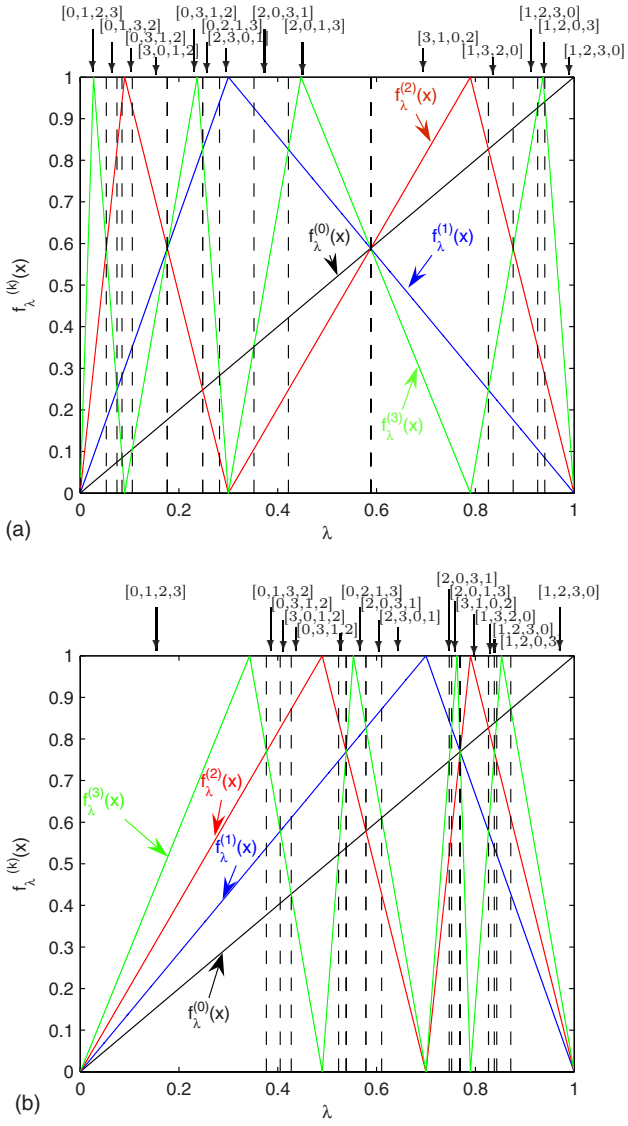


FIG. 3. (Color online) The first four iterations of $f(x)$ and the corresponding order patterns of length 4 for the skew tent map, i.e., $f_\lambda^{(k)}(x)$ for $k=0,1,2,3$. (a) $\lambda=0.3$, (b) $\lambda=0.7$.

## IV. GRAY CODES AND UNIMODAL MAPS

Symbolic dynamics has been thoroughly studied in the context of unimodal maps since the seminal contribution of Metropolis *et al.*[1] In Ref. 3 Gray codes were used as a more intuitive way of understanding and applying the ideas of Ref. 1. The connection between both approaches can be mathematically established with the aid of results in Refs. 1, 2, and 11 as pointed out in Ref. 12. In this section, we address the ordinal structure of Gray codes.

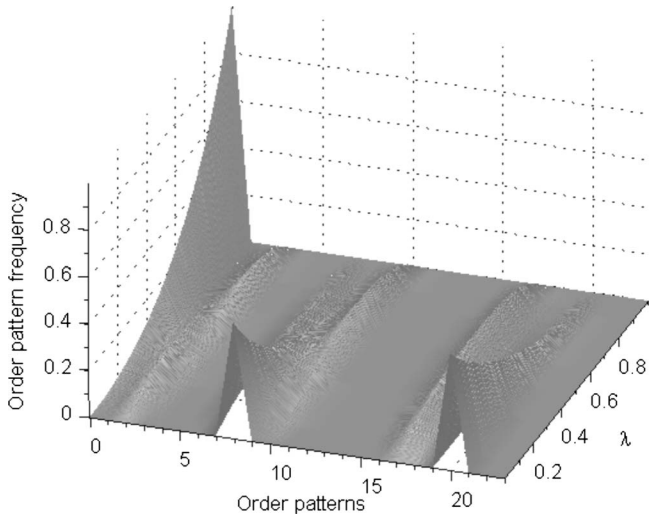


FIG. 4. Relative frequencies of the order patterns of length $L=4$ realized by the skew tent map.

For a unimodal map $f$ defined on the interval $I=[a,b]$, any finite orbit $\{f^n(x):0\leq n\leq N-1\}$ can be transformed into a binary sequence $G_N(f,x)=g(f^0(x))g(f^1(x))\cdots g(f^{N-1}(x))$, where $g:[a,b]\to\{0,1\}$ is the step function

$$g(x)=\begin{cases}0 & \text{if } x<x_c,\\ 1 & \text{if } x\geq x_c.\end{cases} \tag{7}$$

As $x$ increases from the left end point $a$ to the right end point $b$, the interval $I$ can be partitioned into $2^N$ subintervals $I_j^{(N)}$ and $1\leq j\leq 2^N$, each subinterval containing those $x\in I$ whose orbits have resulted into a given binary sequence $G_N(f,x)$. That is (i) $I_j^{(N)}\cap I_j^{(N)}=\emptyset$ for $j\neq k$, (ii) $I=I_1^{(N)}\cup I_2^{(N)}\cdots\cup I_{2^N}^{(N)}$, and (iii) the binary sequences $G_N(f,x)$ obtained for each $x\in I_j^{(N)}$ are the same. Moreover, the sequences $G_N(f,x_1)$ for $x_1\in I_j^{(N)}$ and $G_N(f,x_2)$ for $x_2\in I_{j+1}^{(N)}$, $1\leq j\leq 2^N-1$ differ only in 1 bit. Therefore, if we label the $2^N$ subintervals $I_j^{(N)}$ with the $2^N$ sequences $G_N(f,x)$, then the labels of contiguous subintervals will have only 1 bit flipped.

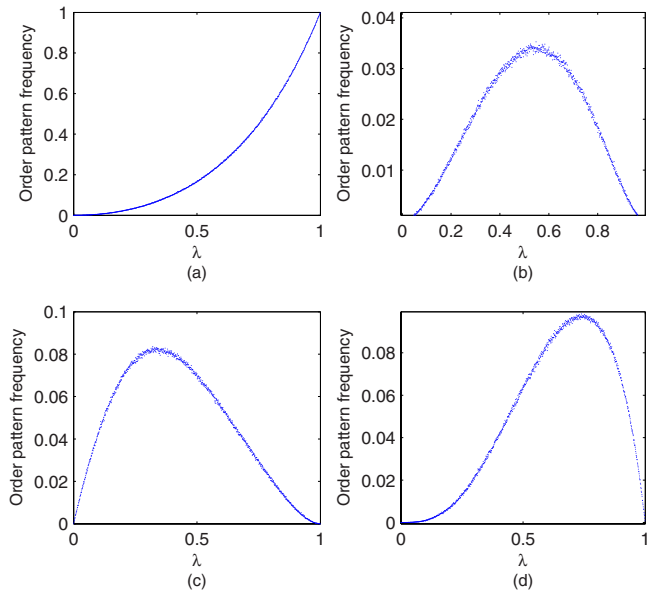


FIG. 5. (Color online) Order pattern frequency for the skew tent map and $L=4$: (a) order pattern 0, (b) order pattern 1, (c) order pattern 2, and (d) order pattern 3.

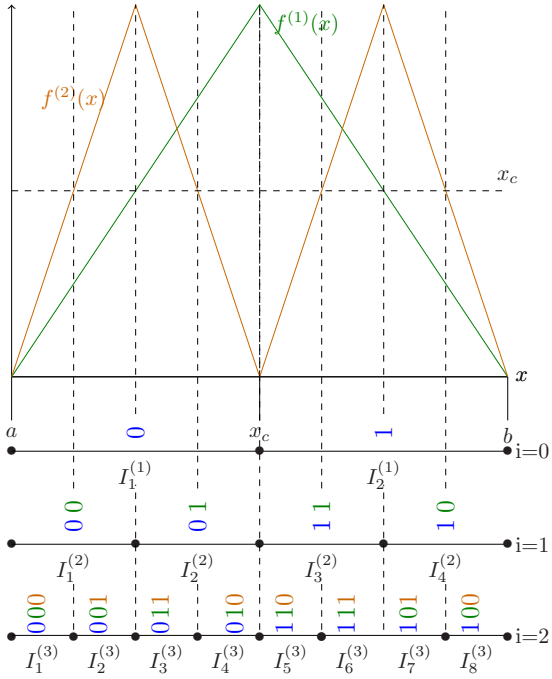


FIG. 6. (Color online) Symbolic intervals for different iterations of the skew tent map for $\lambda=1/2$.

For the sake of illustration, let us consider the skew tent map with $\lambda=1/2$. In Fig. 6, the division of $I=[0,1]$ into the subintervals $I_j^{(N)}$, each labeled with the corresponding binary sequence of length $N$, is shown for $N=1,2,3$. The separation points of the subintervals $I_j^{(N)}$ are the solutions of the equation

$$f_{1/2}^{n-1}(x)=\tfrac{1}{2},\quad 1\leq n\leq N. \tag{8}$$

If, furthermore, $\mathcal{G}_N$ is the set of all binary sequences of length $N$ produced by a map $f\in\mathcal{F}$, then it is possible to endow $\mathcal{G}_N$ with a linear order as follows. Given $G_N(f,x_1)\neq G_N(f,x_2)$, let $i$ be the first index such that $g(f^i(x_1))\neq g(f^i(x_2))$. Depending on the value of $i$, we distinguish three cases:

(1) If $i=0$ then $G_N(f,x_1)<G_N(f,x_2)$ if and only if $g(x_1)<g(x_2)$.
(2) If $i>0$ and $G_i(f,x_1)=G_i(f,x_2)$ contains an even number of 1, then $G_N(f,x_1)<G_N(f,x_2)$ if and only if $g(f^i(x_1))<g(f^i(x_2))$.
(3) If $i>0$ and $G_i(f,x_1)=G_i(f,x_2)$ contains an odd number of 1, then $G_N(f,x_1)<G_N(f,x_2)$ if and only if $g(f^i(x_1))>g(f^i(x_2))$.

Gray codes are well known in the context of communication theory. The Gray codes of length 3 are shown in Table II. The main characteristic of the Gray codes is that two consecutive codes differ in only 1 bit. Moreover, the order of Gray codes is equivalent to the order in $\mathcal{G}_N$ (check Table II for $N=3$). As a consequence, any binary sequence $G_N(f,x)$ can be interpreted as a Gray code of length $N$ (Ref. 3) and will be called a Gray code hereafter. Finally, the order of the

TABLE II. Correspondence between Gray codes and binary codes for 3 bits.

| Rank | Binary code | Gray code |
|---|---|---|
| 0 | 000 | 000 |
| 1 | 001 | 001 |
| 2 | 010 | 011 |
| 3 | 011 | 010 |
| 4 | 100 | 110 |
| 5 | 101 | 111 |
| 6 | 110 | 101 |
| 7 | 111 | 100 |

Gray codes derived from any unimodal map belonging to $\mathcal{F}$ is directly linked to the order in $\mathbb{R}$ of the points $x \in I$. Indeed, (i) $x_1 < x_2$ implies $G_N(f,x_1) \leq G_N(f,x_2)$ for $N \geq 0$ and (ii) $G_N(f,x_1) < G_N(f,x_2)$ for some $N \geq 1$ implies $x_1 < x_2$ (Ref. 11, Lemma 4.1). This is illustrated in Fig. 6.

## V. GRAY CODES AND ORDER PATTERNS FOR UNIMODAL MAPS

In this section the analysis focuses on the parametric unimodal maps of the subclass $\mathcal{F}_1$ or $\mathcal{F}_2$. In Sec. III we elaborated on the dependence of the order patterns allowed for those maps with respect to the control parameter. Specifically, we estimated the probabilities of order 4-patterns by their relative frequencies in orbits of the logistic map (Fig. 2) and of the skew tent map (see Fig. 4) with different parameter settings. Our next goal is to reproduce the same dependencies not from the exact values of the orbit point ("sharp orbit") but from the binary sequence built, as explained in Sec. IV (coarse-grained orbit). As discussed in Sec. IV, the definition domain $I$ of $f \in \mathcal{F}$ splits in $2^N$ subintervals when Gray codes of length $N$ are considered. We show next that the order patterns of $f$ can also be obtained comparing Gray codes obtained from its orbits.

Let $G_M(f,x) = g_0 g_1 \cdots g_{M-1}$, $g_i \in \{0,1\}$ be the Gray code of length $M$ of $x \in I$. Since the Gray codes together with the points $x \in I$ are linearly ordered and, as we saw, their order relations are equivalent, we can expect to obtain useful information about the order patterns realized by the sharp orbit $\mathcal{O}_f(x)$ from the order patterns realized by the coarse-grained orbit $G_M(f,x)$, $M \geq 2$. The procedure is as follows:

(1) Divide the Gray code of length $M$, $G_M(f,x)$ into $M-N+1$ Gray codes of length $N<M$ using a sliding window of length $N$. Thus, the first Gray code derived from $G_M(f,x)$ is $G^0 = g_0 g_1 \cdots g_{N-1} = G_N(f,x)$, the second Gray code is $G^1 = g_1 g_2 \cdots g_N = G_N(f,f(x)), \ldots$, and the $(M-N+1)$th Gray code is $G^{M-N} = g_{M-N} g_{M-N+1} \cdots g_{M-1} = G_N(f,f^{M-N}(x))$.

(2) For $i = 0,1,\ldots,M-N-L+1$, build groups of $L$ consecutive Gray codes $G^i G^{i+1} \cdots G^{i+L-1}$. The $i$th group defines then the order $L$-pattern $\pi = \pi(i) = [\pi_0, \pi_1, \ldots, \pi_{L-1}]$ if

$$G^{i+\pi_0} < G^{i+\pi_1} < \cdots < G^{i+\pi_{L-1}}.$$

The order patterns derived using Gray codes need not have, in general, similar λ-DFs to those derived from the sharp orbits. Indeed, order patterns defined by Gray codes of length $N$ are built upon the comparison of subintervals $I_j^{(N)} \subset I$ (see Sec. IV) rather than comparing points of $I$. The width of the intervals $I_j^{(N)}$ decreases as the length $N$ of the sliding window increases in such a way that when $N \to \infty$, each one of those intervals converges to a single real number. As a result, the error in the calculation of the order patterns from Gray codes is expected to reduce as $N$ increases. In the context of finite-precision computation, the minimum value of $N$ necessary to get a reliable approximation of the λ-DF of an order pattern is related to the precision of the arithmetic used. Again, this quantization error decreases as $N$ increases and, consequently, a large value of $N$ may be necessary to assure a good approximation of the λ-DF.

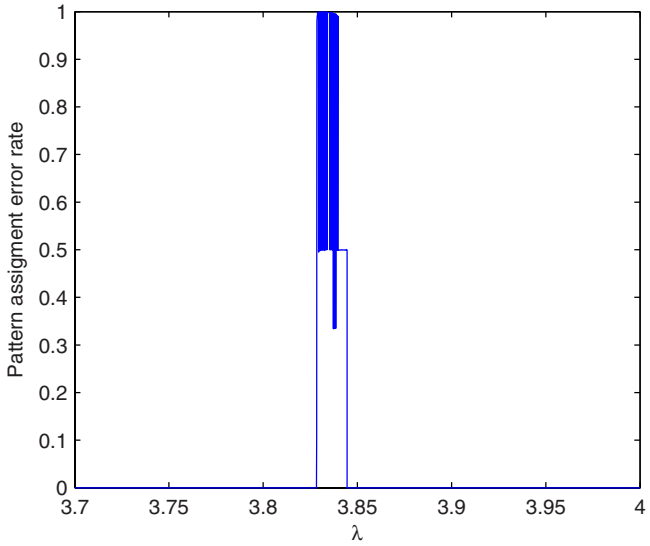


FIG. 7. (Color online) Error rate for the pattern assignment based on Gray codes with respect to the one based on the orbit of the logistic map. The length of order patterns is $L=4$, the length of the considered Gray codes is $N=100$, and the number of samples is 10 104. The perfect estimation of the PDF of the order patterns of the logistic map is only possible for those values of λ leading to ergodic behavior.

Another source of divergences between λ-DFs and their numerical estimation via finite-length Gray codes is nonergodicity or even poor ergodicity. As a matter of fact, remember that the estimation of the probability $\mu(P_\pi)$ by the relative frequency of $\pi \in \mathcal{S}_L$ in finite orbits of a $\mu$-preserving map hinges on the ergodic theorem. If, furthermore, the convergence of relative frequencies to probabilities in the orbits of an ergodic map with respect to $\mu$ is very slow, a good estimation would require exceedingly long sequences—this is what we mean by "poor ergodicity." These errors are shown in Figs. 7 and 8 for the logistic and the skew tent maps, respectively, with $\pi = [0,1,2,3]$, $M = 10\,104$, and $N = 100$. In the first case, the value of λ lies within the period-three window of the logistic map. In the second case, poor ergodicity is expected for values of λ close to 0 and 1. The asymmetry in the error distribution is due to the fact that for $\lambda \simeq 1$, the tent map looks like the identity in most of $I = [0,1]$, hence $P_{[0,1,2,3]}$ covers most of $I$. This makes [0,1,2,3] to be the most frequent order 4-pattern even when its frequency is calculated using Gray codes. Comparison of Figs. 5 and 9 illustrates the accuracy of the Gray code-based

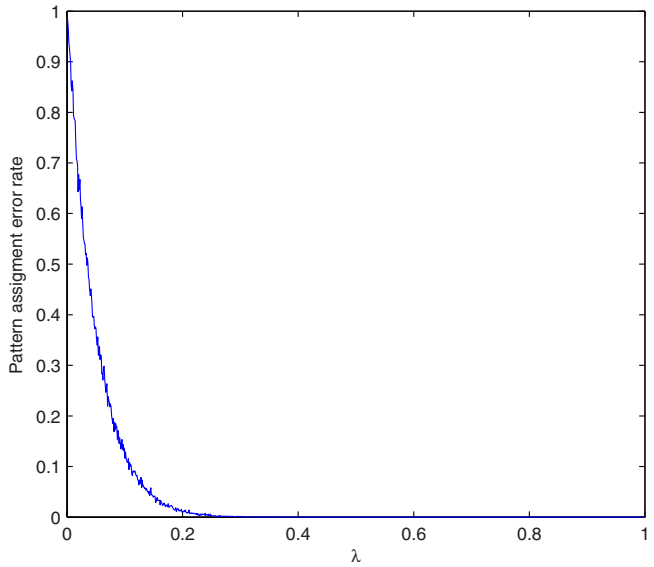


FIG. 8. (Color online) Error rate for the pattern assignment based on Gray codes with respect to the one based on the orbit of the skew tent map. The length of order patterns is $L=4$, the length of the considered Gray codes is $N=100$, and the number of samples is 10 104. A value of the control parameter above 0.2 guarantees a perfect estimation of the PDF of the order patterns of the skew tent map.

method for the first four order 4-patterns (see Table II) of the skew tent map.

## VI. ESTIMATION OF THE CONTROL PARAMETER FOR UNIMODAL MAPS WITH CRITICAL POINT DEPENDING ON THE CONTROL PARAMETER

The main characteristic of the maps in $\mathcal{F}_2$ is that the control parameter λ determines the value of the critical point. Furthermore, from our discussion above, we expect that the relation between the control parameter and the allowed order

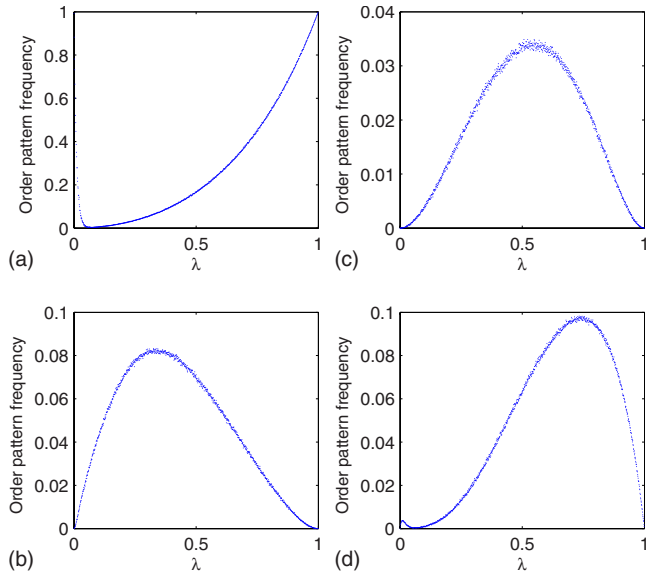


FIG. 9. (Color online) Relative frequency of order patterns of the skew tent map using Gray codes when $L=4$, $N=100$ and the sequences are 10 104 bit long: (a) order pattern 0, (b) order pattern 1, (c) order pattern 2, and (d) order pattern 3.

patterns of the corresponding dynamics is specially simple for the pattern $\pi=[0,1,\ldots,L-1]$. Clearly, if the λ-DF of this pattern is one-to-one, then λ can be pinpointed from that distribution function; otherwise, the possible values of λ can be reduced to a few candidates, which can be also acceptable in applications such as cryptanalysis. In turn, λ-DFs can be approximated via Gray codes without previous knowledge on the critical point of the map. The bottom line is that the control parameter of a map in $\mathcal{F}_2$ can be estimated from their coarse-grained orbits (in form of Gray codes). The specifics depend on the map.

In more general terms, let $f_\lambda \in \mathcal{F}_2$ and suppose that each $f_\lambda$ is ergodic for $\lambda \in J$ with the same invariant measure $\mu$. Furthermore, assume for the time being that $f_\lambda(a)=a$ and $f_\lambda(x)>x$ on a nonempty vicinity of $a$. Let $(a,c)$ be the maximal interval in $(a,x_c)$ such that $x<f_\lambda(x)$. We claim that the interval

$$I_L^\lambda = (a,c) \cap f_\lambda^{-1}(a,c) \cap \ \ldots \ \cap f_\lambda^{-(L-1)}(a,c) \tag{9}$$

coincides with $P_{[0,1,\ldots,L-1]}$. Indeed, if $x \in I_L^\lambda$, then $f^i(x) \in (a,c)$ for $0 \le i \le L-1$ and

$$x < f_\lambda(x) \Rightarrow f_\lambda(x) < f_\lambda^2(x) \Rightarrow \cdots \Rightarrow f_\lambda^{L-2}(x) < f_\lambda^{L-1}(x).$$

Hence, $I_L^\lambda \subset P_{[0,1,\ldots,L-1]}$. Conversely, if $x \in P_{[0,1,\ldots,L-1]}$, i.e.,

$$x < f_\lambda(x) < f_\lambda^2(x) < \cdots < f_\lambda^{L-1}(x),$$

then $f^i(x) \in (a,c)$ for $0 \le i \le L-1$. Thus, $P_{[0,1,\ldots,L-1]} \subset I_L^\lambda$. This proves

$$P_{[0,1,\ldots,L-1]} = I_L^\lambda.$$

If, otherwise, $f_\lambda(a)=a$ but $f_\lambda(x)<x$ on a nonempty vicinity of $a$, let $(a,c)$ be the maximal interval in $(a,x_c)$ such that $f_\lambda(x)<x$. In this case, a similar reasoning (changing the direction of the inequalities) shows that $I_L^\lambda = P_{[L-1,L-2,\ldots,1,0]}$. For definiteness we shall consider henceforth only the interval $P_{[0,1,\ldots,L-1]}$ (similar arguments apply *mutatis mutandis* to the interval $P_{[L-1,\ldots,0,1]}$). Because of ergodicity, the relative frequency at which a typical trajectory visits $I_L^\lambda$ is $\mu(I_L^\lambda)$. If $\mu(I_L^\lambda)$ happens to be different for each interval $I_L^\lambda$, then $\mu(I_L^\lambda)$ can be used to determine or estimate the control parameter λ. In this case, the frequency of the order pattern $[0,1,\ldots,L-1]$ in an orbit $\mathcal{O}_{f_\lambda}(x)$ is just the number of times that $f_\lambda^{i+j}(x) \in (a,c)$ for $i \in \mathbb{N}_0$ and $j=0,1,\ldots,L-1$.

As an example, consider the skew tent map again. For this map, the interval $P_{[0,1,\ldots,L-1]}$, i.e., the set of points $x \in [0,1]$ of type $[0,1,\ldots,L-1]$, is determined by the leftmost intersection of the iterates $f_\lambda^{L-2}$ and $f_\lambda^{L-1}$, where

$$f_\lambda^n(x) = \begin{cases} x/\lambda^n & \text{if } 0 \le x \le \lambda^n, \\ (\lambda^{n-1}-x)/\lambda^{n-1}(1-\lambda) & \text{if } \lambda^n \le x \le \lambda^{n-1}. \end{cases} \tag{10}$$

Hence $P_{[0,1,\ldots,L-1]}=(0,\phi_L(\lambda))$ with

$$\phi_L(\lambda) = \frac{\lambda^{L-2}}{2-\lambda}. \tag{11}$$

Since this function is one-to-one in the interval $0 \le \lambda \le 1$ for $L \ge 2$ with $\phi_2(0)=1/2$, $\phi_{L\ge 3}(0)=0$, and $\phi_{L\ge 2}(1)=1$, it al-

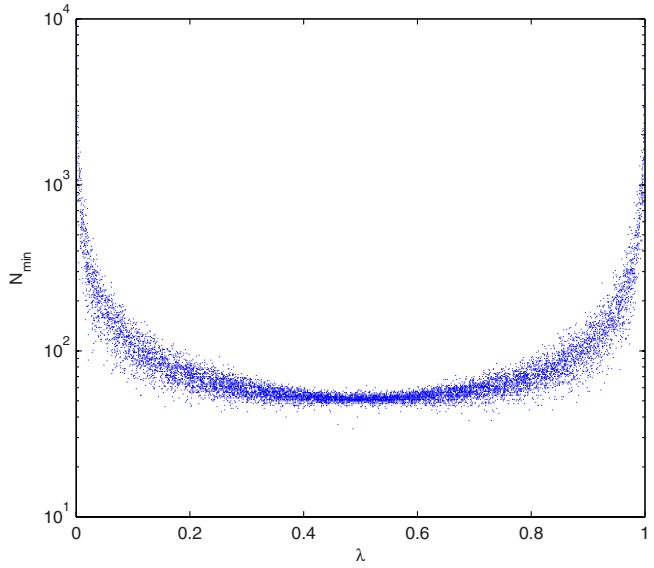


FIG. 10. (Color online) Minimum width of the sliding window necessary for the reconstruction of the PDF of the order patterns from the symbolic sequences of the skew tent map.

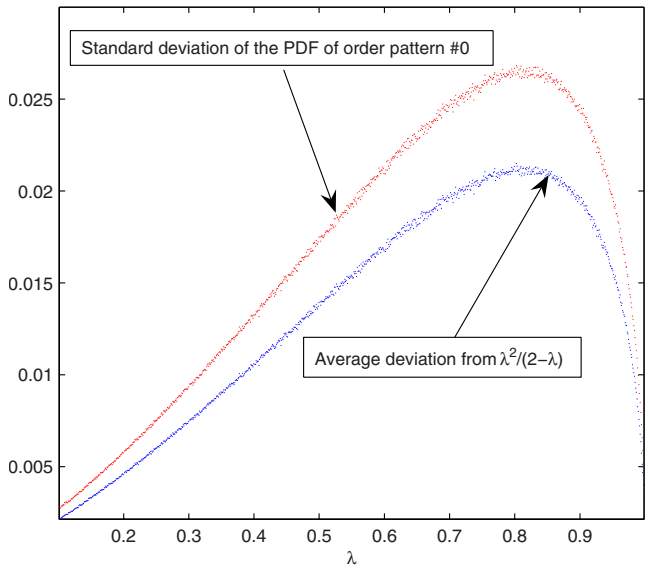


FIG. 11. (Color online) Average deviation and standard deviation in the estimation of the PDF of the order pattern 0 for the skew tent map with respect to $\lambda^2/(2-\lambda)$.

lows to estimate λ by estimating $\phi_L(\lambda)$—the length of $P_{[0,1,\ldots,L-1]}$. Now, from the equation

$$\frac{d}{d\lambda}\phi_L(\lambda) = \frac{\lambda^{L-3}}{(2-\lambda)^2}[2(L-2)-(L-3)\lambda] = \begin{cases} 0, & \text{if } \lambda = 0, \\ L-1, & \text{if } \lambda = 1, \end{cases} \tag{12}$$

it follows that $\phi_L(\lambda)$ is a ∪-convex function on $0 \le \lambda \le 1$ for $L \ge 2$ that converges to 0 on $0 \le \lambda < 1$ as $L \to \infty$. Therefore, the higher the $L$, the worse $\phi_L(\lambda)$ discriminates different values of λ. Consequently, $L=3,4$ are the best choices for a quality estimation of λ.

On the other hand, the ergodicity of the skew tent map permits to estimate the length of $P_{[0,1,\ldots,L-1]}$ by estimating the relative frequency of $\pi=[0,1,\ldots,L-1]$ in a typical sharp orbit of the map, or, as we intend, in a typical coarse-grained orbit. In the latter case, the choice for the parameter $N$, the width of the sliding window down the Gray codes (Sec. V), must be also analyzed. The minimum value of $N$ to get a good reconstruction of the λ-DF of the order patterns $N_{\min}$ depends on the precision of the arithmetic used, but it also depends on the Lyapunov exponent of the map. If floating point double-precision arithmetic is implemented, then $N_{\min}$ can be determined as a function of λ by comparing pairs of symbolic sequences generated from the same initial condition with control parameters $\lambda_1$ and $\lambda_2$, such that $|\lambda_2-\lambda_1|$ equals the spacing of floating point numbers. Figure 10 shows the dependence of $N_{\min}$ with respect to λ.

Summing up, the estimation of the control parameter $\lambda \in (0,1)$ of the skew tent map $f_\lambda$, Eq. (5), can be done by counting and normalizing the occurrences of the order pattern $[0,1,\ldots,L-1]$ ideally for $L=3$ or 4 in a statistically significant sample of orbit segments of $f_\lambda$. This follows from the following properties: (i) $f_\lambda$ is ergodic for all λ and (ii) the $f_\lambda$-invariant measure of $P_{[0,1,\ldots,L-1]}$ (in this case, the length of the interval $P_{[0,1,\ldots,L-1]}$) depends bijectively on λ. In a practical context though, finite precision machines are used and this entails, in general, numerical degradation, meaning that the computed orbits, whether of chaotic or nonchaotic maps, depart from the real ones. In the case of a very long orbit of a chaotic map, the deviation of the numerical simulation (locally measured by the Lyapunov exponent of the map) will be severe; in such cases, it is preferable to have many shorter orbits instead. Even worse, all orbits computed with finite precision are eventually periodic. This distortion of the dynamics due to finite numerical precision and dependence on initial conditions implies the general impossibility of obtaining orbits and invariant measures in an exact way. As a matter of fact, all this carries over to symbolic dynamics.

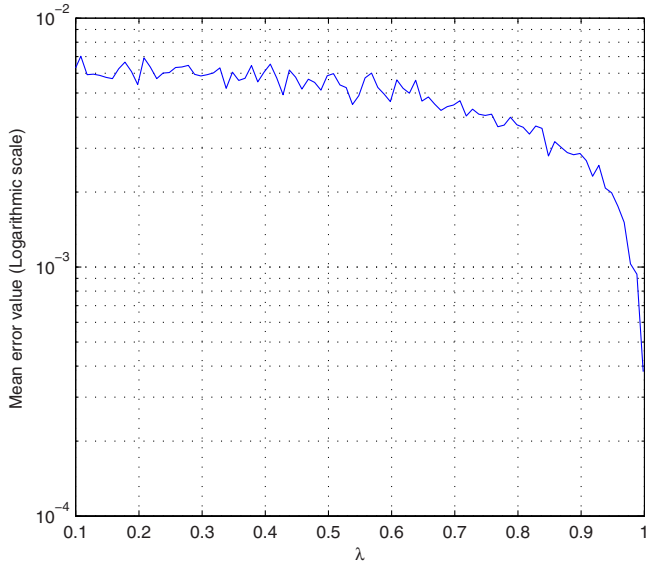


FIG. 12. (Color online) Mean error value in the estimation of the control parameter of the skew tent map.

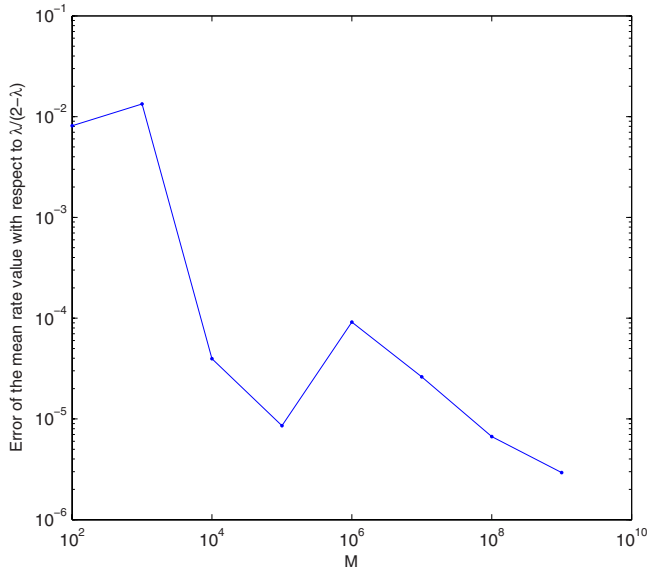


FIG. 13. (Color online) Dependency of the error in the estimation of the rate of occurrences of the order pattern 0 with respect to the length of the orbits.

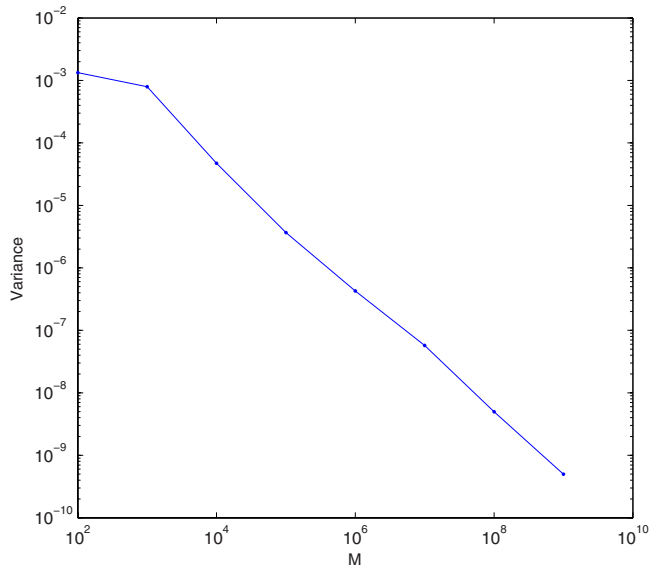


FIG. 14. (Color online) Variance of the error in the estimation of the rate of occurrences of the order pattern 0 with respect to the length of the orbits.

To verify this issue in the case of coarse-grained orbits, a sample of Gray codes of the skew tent map, each one with the same length but with a different initial condition, was generated for every value of λ. The underlying sharp orbits were computed with double precision floating point arithmetic. From this sample of Gray codes, the corresponding λ-DFs of the order patterns of length $L=4$ were obtained. The λ-DF of the order pattern [0,1,2,3] (0 for short) was calculated as the mean value of the λ-DFs obtained from the various initial conditions. This average value is compared to the exact λ-DF, $\phi_4(\lambda)=\lambda^2/(2-\lambda)$, in Fig. 11 along with the corresponding standard deviation.

Figure 11 spells out that in the context of finite precision computation, the perfect recovery of the control parameter value using the λ-DF of order pattern 0 is not feasible in general if one can only resort to Gray codes. However, it is possible to locate λ up to an uncertainty interval. The width of this interval can be upper bounded by the standard deviation of the λ-DF of the order pattern 0 since, according to Fig. 11, it is bigger than the average error in the estimation of $\phi_4(\lambda)$ for every value of λ. Therefore, the estimation of the control parameter comprises two stages:

(1) An estimation of λ is performed by dividing the given Gray code $\{g_i\}_{i=0}^{M-1}$, $g_i \in \{0,1\}$ into a large enough set of disjoint subsequences of length $N \gg 4$, say, $\{g_{kN+i}\}_{i=0}^{N-1}$ for $k=0,1,\ldots,K=\lfloor M-N \rfloor-1$. For each such binary subsequence, a value of $\phi_4(\lambda)$ is then computed as the relative frequency of the order pattern [0,1,2,3] using, of course, the Gray ordering (Sec. IV). Let $\bar{x}$ be the mean value of the resulting $\phi_4(\lambda)$. From $\phi_4(\lambda)=\lambda^2/(2-\lambda)$, Eq. (6), it follows that the control parameter is estimated as

$$\hat{\lambda} = \phi_4^{-1}(\bar{x}) = \frac{-\bar{x}+\sqrt{\bar{x}^2+8\bar{x}}}{2}. \tag{13}$$

(2) If $\sigma$ is the standard deviation of the $\phi_4(\lambda)$ sampling, then there exists a high probability that λ is in the interval

$$(\phi_4^{-1}(\bar{x}-\sigma),\ \phi_4^{-1}(\bar{x}+\sigma)). \tag{14}$$

The specifics of this procedure refer to the skew tent map, but the general strategy is the same once a bijective λ-DF of an order pattern is exactly known. The one-parameter tent map [Eq. (5)] is just a specimen of a more general family: unimodal, piecewise linear expanding Markov transformations. Moreover, each topologically mixing transformation in this family (i.e., some power of its transition matrix is strictly positive) has a unique ergodic invariant measure,[13] which furthermore is absolutely continuous with respect to the Lebesgue measure. This measure can be calculated or numerically estimated by a variety of methods (Perron–Frobenius operator, Ulam's method, or just a computation of long time averages). For the purpose envisaged in this paper, an exact knowledge on the invariant measures is though not necessary, since the λ-DF of order patterns can be calculated with numerical simulations. The important features are (i) ergodicity, so that the statistical properties of the orbits are independent of the initial condition, and (ii) the absolute continuity of the (unique) invariant measure, which guarantees that it is accessible to numerical methods. In sum, piecewise linear expanding Markov transformations are a handy source of maps with the properties needed. Furthermore, given one such map, one can produce a one-parameter family of them by stretching and compressing the intervals of the corresponding Markov partition while keeping the maps expanding. A different question is whether the resulting λ-DF will provide a one-to-one relation between the relative frequency of some order pattern and the parameter λ. For cryptographic applications, a few-to-one relation may be sufficient.

In order to establish the accuracy of the procedure, some numerical simulations with the skew tent map were done. For every value of the control parameter, a group of 200

different initial conditions was used in the generation of the corresponding Gray codes. For each of these binary sequences, the control parameter λ was estimated as just explained. The mean error of the estimation is shown in Fig. 12. The average error lies always above $10^{-4}$ and can only be reduced by the implementation of the procedure with extended-precision arithmetic libraries.

To prove this claim, the case of the symmetric tent map, i.e., the skew tent map for $\lambda=1/2$, will be now considered. For the symmetric tent map the arithmetic is exact. Indeed, if $0.x_0x_1\cdots x_M$, $x_i\in\{0,1\}$ is the expansion to base 2 of $x\in[0,1]$, i.e.,

$$x=\sum_{n=0}^{M}\frac{x_n}{2^{n+1}} \tag{15}$$

(numbers with finite binary expansions are called dyadic rationals), then the action of the symmetric tent map amounts to a 0-bit dependent left shift, to wit:

$$f_{1/2}(0.x_0x_1\cdots x_n\cdots x_{M-1}x_M)=\begin{cases}0.x_1x_2\cdots x_{n+1}\cdots x_M0 & \text{if } x_0=0,\\ 0.x_1^*x_2^*\cdots x_{n+1}^*\cdots x_M^*0 & \text{if } x_0=1,\end{cases} \tag{16}$$

where $x_n^*=1-x_n$. Therefore, if $x\in[0,1]$ is represented with $M$ bits and $x_M=1$, the orbit of $x$ collapses to 0 after $M$ iterations of $f_{1/2}$, so $M$ can be considered the effective length of the orbits to be used in an estimation of $\lambda=1/2$. For $L=3$, the relative frequency of the order pattern 0 ([0,1,2] in this case) was determined for a large set of random initial conditions $x$ and increasing orbit lengths $M$. The convergence in average of this relative frequency to $\phi_3(1/2)=1/3$ [see Eq. (11)] as $M$ increases is confirmed by Fig. 13. At the same time, the variance of the estimation steadily reduces with $M$, as shown in Fig. 14. In other words, a higher precision of the arithmetic used in orbit generation and greater samples for the subsequent control parameter estimation clearly improves the results.

We conclude that the inaccuracies exposed above in our method to recover the control parameter of maps of $\mathcal{F}_2$ based on the order patterns of their coarse-grained orbits (specifically in form of Gray codes) are due to the shortcomings of finite precision arithmetic and finite statistical sampling but are not inherent to the method, as proved with the symmetric tent map.

## VII. CONCLUSIONS

In this paper it was shown how to rebuild the λ-DFs of order patterns from Gray codes, the scope being the estimation of the parameter λ. Gray codes are the 0-1 sequences that result from the symbolic dynamics of unimodal maps with respect to the left-right partition of the state space introduced by the critical point. We have analyzed the λ-DFs of the order patterns of two unimodal parametric maps: the logistic map (as representative of the subclass $\mathcal{F}_1$) and the skew tent map (as representative of the subclass $\mathcal{F}_2$). In the case of the logistic map, it turns out that this technique can hardly deliver on account of the complex and many-to-one relation between λ and those λ-DFs. On the contrary, this relationship is simple, one-to-one, and analytically known for $\pi=[0,1,\ldots,L-1]$ in the case of the skew tent map. Our method improves previous proposals for parameter estimation from symbolic sequences of unimodal maps in that a knowledge on the critical point value is not needed. The most important consequence of this paper is that symbolic sequences of unimodal maps cannot be used as key streams of stream ciphers. Indeed, the work described in Ref. 14 along with our method reveals a critical vulnerability in encryption systems such as the one introduced in Ref. 15. However, our method demands high computational precision and large amounts of data. In this regard, we recommend the use of extended-precision libraries for good estimations. In the ideal case of arbitrarily high precision, the estimated value of the control parameter is arbitrarily close to the real one.

## ACKNOWLEDGMENTS

The authors are very grateful to the reviewers for their constructive criticism. The work described in this paper was supported by Ministerio de Educación y Ciencia of Spain, Research Grant No. SEG2004-02418, CDTI, Ministerio de Industria, Turismo y Comercio of Spain in collaboration with Telefónica I+D, Project SEGUR@ with reference CENIT-2007 2004, CDTI, Ministerio de Industria, Turismo y Comercio of Spain in collaboration with SAC, Project HESPERIA (CENIT 2006–2009), and Ministerio de Ciencia e Innovación of Spain, Project CUCO (Grant No. MTM2008-02194).

[1] N. Metropolis, M. Stein, and P. Stein, J. Comb. Theory, Ser. A **15**, 25 (1973).
[2] L. Wang and N. D. Kazarinoff, J. Comb. Theory, Ser. A **46**, 39 (1987).
[3] G. Alvarez, M. Romera, G. Pastor, and F. Montoya, Electron. Lett. **34**, 1304 (1998).
[4] J. M. Amigó, L. Kocarev, and J. Szczepanski, Phys. Lett. A **355**, 27 (2006).
[5] J. M. Amigó, S. Zambrano, and M. A. F. Sanjuán, Europhys. Lett. **83**, 60005 (2008).
[6] D. L. Kreher and D. R. Stinson, *Combinatorial Algorithms: Generation, Enumeration and Search* (CRC, Boca Raton, Florida, 1998).
[7] J. M. Amigó, S. Elizalde, and M. B. Kennel, J. Comb. Theory, Ser. A **115**, 485 (2008).
[8] P. Walters, *An Introduction to Ergodic Theory*, Graduate Texts in Mathematics Vol. 79 (Springer-Verlag, Berlin, 1982).
[9] J. M. Amigó, S. Zambrano, and M. A. F. Sanjuán, Europhys. Lett. **79**, 50001 (2007).
[10] L. Billings and E. M. Bollt, Chaos, Solitons Fractals **12**, 365 (2001).
[11] W. Beyer, R. Mauldin, and P. Stein, J. Math. Anal. Appl. **115**, 305 (1986).
[12] T. Cusick, Electron. Lett. **35**, 468 (1999).
[13] A. Lasota and J. A. Yorke, Trans. Am. Math. Soc. **186**, 481 (1973).
[14] X. Wu, H. Hu, and B. Zhang, Chaos, Solitons Fractals **22**, 359 (2004).
[15] A. P. Kurian and S. Puthusserypady, Signal Processing **88**, 2442 (2008).